CONTENTS

Abstract……..……..... i

List of Figures……..……...……..……...…….……iii

1. Introduction……..……..……..……..……..……..……..1

2. Aim of Study……...……..……..……..……...3

3. Objectives of study……...……..……..……..……...4

4. Study background (Literature review) and Significance of Study……..……..4

5. National status ……..……..……..……..……...…….17

6. International status ……..……..……..……...……..18

7. Study Hypothesis ……..……...……..……..……..19

8. Study Rationale……..……...……..……..……..19

9. Research Methodology ……...……..……...……...……20

10. Statistical Plan……..……...……..……..……..23

11. Limitations of the Study……..……...……..……..……..24

12. Expected Outcomes of Study……..……..……..……..24

13. Conclusion……..……...……..……..……..25

14. Bibliography……..……...……..……..…….26

LIST OF FIGURES

Fig. 1.1. Taxonomy of Text Summarization ……...…2

Fig. 1.2. The basic workflow of Extractive & Abstractive summarization……..2

Fig. 1.3. Proposed Methodology of Legal Document Summarization……...…21

Design Framework for Legal Document Summarization

1. Introduction:

Natural Language Processing (NLP) is a very important field of artificial intelligence that works on teaching machines how to perceive, interpret, and create human language. It includes a lot of different tasks, including as sentiment analysis, machine translation, Named Entity Recognition (NER), and text summarization. Text summarization is one of the most important ways to shorten and make sense of enormous amounts of text data while keeping the main idea and information [2]. Processing legal text data is hard because it is complex, irregular, and full of domain-specific jargon. Legal documents such case laws, judgments, and contracts are usually long and complicated, which makes it hard for even experienced lawyers to find the information they need. Also, finding and putting together relevant information from many legal papers takes a lot of time and often requires qualified experts to do it by hand. This not only makes things harder to think about, but it also costs a lot of money and takes a long time to get things done in court. In this situation, automatic summarization has become an important part of legal NLP, with the goal of making things easier to use and more efficient. AI-powered summarization tools aim to provide concise, practical summaries that support case analysis, compliance, and legal research. [47]

The performance of legal text summarization platforms has seen substantial enhancements due to recent innovations in Explainable AI (XAI) and hybrid summarization methods. The finest characteristics of both extractive summarizing which maintains the original wordings of the law and abstractive summarization which paraphrases and shortens the text to be more readable and understandable are blended in hybrid methods. Collectively, these methods generate summaries that are true yet also context-conscious and meaningful. Concurrently, XAI methods guarantee that AI systems employed in high-risk legal cases are understandable and transparent, something which is very important. Due to such developments, this study proposes an intelligent system for summarizing court documents and utilizing explainable AI, hybrid summarization, legal keyword extraction, and domain-specific language models.

The system is meant to make accurate summaries based only on the legal document's content. A multilingual module also makes the framework useful for legal papers that span borders. When it comes to summarizing legal documents, it's important to know about the numerous sorts of summary approaches. There are many ways to group text summarization, such as by output type, input source, context, and goal. Figure 1.1 shows that there are different ways to summarize information based on whether the output is extractive or abstractive, whether the input is one document or many, whether the summary is meant to be general or specific to the user, and whether

the goal is to be indicative, informative, or critical. This taxonomy gives you a complete way to choose the best summary methods for various fields, including law. There are two main ways to summarize text: extractive and abstractive. Extractive summarization means picking out important sentences or phrases from the original source to make a summary. The goal is to get the most important information without changing the original text. Abstractive summary, on the other hand, tries to restate and reinterpret the main concepts of a document, as how people might summarize by making new sentences that capture the main ideas. Figure 1.2 shows how extractive and abstractive summarization work in general [1]. These are the basic strategies used in today's AI-driven legal summarization systems. Recent improvements have used hybrid and XAI models to incorporate the best parts of both approaches.

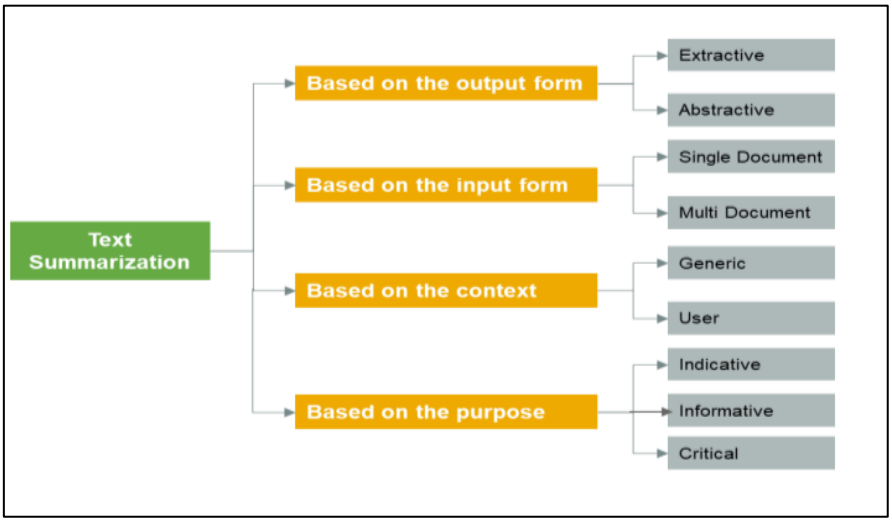

Fig. 1.1. Taxonomy of Text Summarization [47]

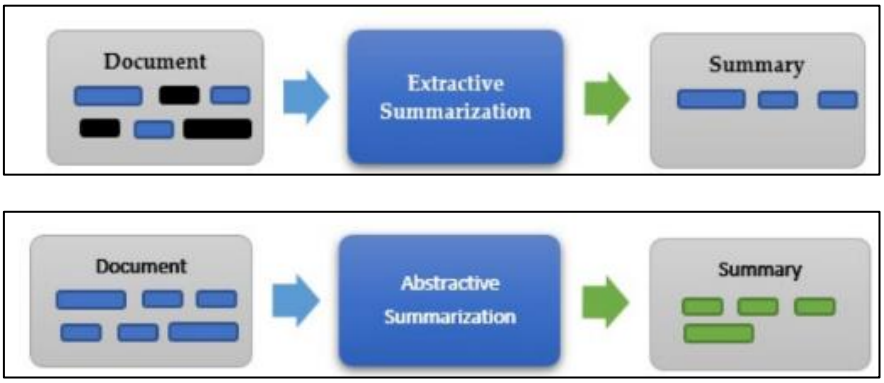

Fig. 1.2. The basic workflow of Extractive & Abstractive summarization [1]

Large language models (LLMs) have shown great promise in many broad natural language processing jobs, but they still have a lot of problems that make them hard to use in the legal field.

These models often make up fake or wrong legal information and usually don't have access to up-to-date or contextually relevant legal materials. This can lead to outputs that aren't legally recognized or don't show the subtleties of reasoning in a certain field. But in legal situations, factual correctness isn't enough; transparency and interpretability are just as important. As AI systems get more complicated, it's important to know how they work on the inside and make sure that users, administrators, and society can trust them. Explainable AI means being able to explain algorithms, models, and decision-making processes in a way that people can understand [1]. This is where XAI comes in. Traditional black-box AI algorithms don't give much or any information about how judgments or summaries are made, which makes it harder to hold people accountable and trust them. XAI approaches deal with these issues by making things clearer, revealing which sections of the original text went into the final summary, and by making things easier to grasp, so that legal professionals can see why AI outputs are what they are. This builds confidence among users and helps AI systems be used more in courts, legal settings, and regulatory settings. XAI also enables it to be simpler to hold individuals accountable by allowing mistakes to be easily discovered and refined as models evolve with time. An XAI-augmented summary may have an explanation such as, "This sentence was added because it is the court's final decision, backed by Section 302 IPC." The synergy created by using the combination of hybrid summarization, legal keyword extraction, and XAI is very strong and results in summarization systems that not only represent accurate and current information, but also easy to understand, comprehendible, and ethically aligned with the requirements of legal professionals. The framework suggested in this research is derived from the integrated approach.

2. Aim of the Study:

The aim of this study is to design an intelligent, legal document processing framework that will implement hybrid summarization approaches (both extractive and abstractive) with specialized language models in the legal domain along with legal keyword extraction, XAI, and support for multilingual capabilities. This framework will be purposefully built to effectively process long and complex legal documents articulately and meaningfully while maintaining high levels of accuracy, and interpretability. The use of hybrid summarization techniques will help alleviate hallucinations by embedding outputs to the original text for grounding and will improve trust through transparent, explainable outputs. By addressing important challenges such as long context understanding, development of a legal domain-specific architecture, and enabling multilingual processing, this study will achieve a robust, exploratory and legally compliant AI system that can deliver legal summarization.

3. Objectives of study:

The objective of this research is to design a legal document summarization framework that combines extractive and abstractive techniques with XAI to generate accurate, coherent, and transparent summaries, and to extend it to support multilingual legal content for broader accessibility.

the study aims:

- To explore and analyse existing legal summarization approaches and datasets for summarization tasks.
- To design a hybrid extractive–abstractive summarization framework integrated with Explainable AI (XAI) techniques.
- To validate the system's performance using quantitative metrics and expert feedback to ensure transparency and interpretability.
- To extend the validated system for multilingual summarization of legal content, enhancing wider accessibility.

4. Study background (Literature Review) and Significance of Study:

Literature Review

The application of artificial intelligence (AI) methods like XAI and NLP has led to quick developments in the legal field. Enhancing legal text summarization, and decision-making support are the goals of these technologies.

The focus of Paper [1] is on using XAI and a hybrid NLP approach to summarize, predict, and analyse rulings from the Turkish Constitutional Court. The study achieves high accuracy in summarization and classification by combining several models, including TF-IDF, word embeddings, and deep learning frameworks. A key contribution is the use of SHAP (SHapley Additive exPlanations), which provides insights into how the model makes decisions. This hybrid model greatly improves the transparency and reliability of automated legal decision systems. The authors note that this approach performs well within the Turkish legal language context and prove its effectiveness through practical experiments. The study recommends future exploration of

multilingual setups and the incorporation of advanced transformers to enhance generalizability and support legal reasoning.

Work in [2] introduces a hybrid transformer-based framework specifically designed for multi-document summarization of Turkish legal documents. The paper discusses the limitations of conventional summarization techniques in handling domain-specific language and introduces a model that integrates BERT-based sentence embeddings with a transformer decoder. The approach enhances coherence, factual accuracy, and redundancy elimination. The authors conducted comprehensive evaluations against standard benchmarks, showing superior performance in summary informativeness and linguistic quality. Identified gaps include the need for domain-specific fine-tuning and inclusion of contextual metadata like legal citations and temporal references. The authors propose future enhancements through reinforcement learning and multilingual adaptation.

Legal Query RAG, as discussed in [3] investigates the use of retrieval-augmented generation for answering legal queries. The authors design a legal assistant system based on GPT-4 with a RAG architecture that fetches case-relevant information in real time. They highlight how the system bridges the gap between traditional retrieval systems and generative language models by embedding domain-specific retrieval for context-rich responses. Key challenges noted include high latency in generation, lack of structured legal reasoning, and issues with factual grounding. The authors propose building domain-adapted evaluation frameworks and incorporating feedback loops with legal experts to fine-tune results, especially for case-law-based jurisdictions.

RAG4DS [4] offers an in-depth analysis of the lifetime, challenges, and potentiality of Retrieval-Augmented Generation in data realms. A single paradigm that bridges data retrieval and generative processes is suggested in the paper, as well as stressing the necessity of interpretability and trustworthiness in high-risk domains such as law. It also points out challenges with document-level understanding, cross-domain retrieval, and safe data integration. In order to address scalability and interoperability issues, the authors recommend applying knowledge graphs and decentralized data architectures. The paper ends with an appeal to the standardized lifecycle management of RAG in data ecosystems in real-world applications, including legal and regulatory frameworks.

In [5], the author examines existing legal NLP techniques using large language models (LLMs). The paper observes that though there has been significant development in the use of LLMs to summarize, translate, and classify law, there is a strong English-centric bias in datasets and

jurisdictions. It highlights the underutilization of explainable AI by legal LLMs and the absence of standardized benchmarks. The paper advises wider evaluation measures that test legal reasoning, factual coherence, and ethical impact. The review supports more multilingual assistance, additional datasets from underrepresented jurisdictions, and greater transparency in model decision-making.

Work in [6] highlights the importance of interpretability and precision in Retrieval-Augmented Generation models employed in legal technology. It offers a critical review of available RAG models as well as their shortcomings, e.g., handling noisy data, narrow domain coverage, and a lack of interpretability. The research classifies RAG limitations into structural, semantic, and ethical aspects, each with proposed recommendations. Future research directions involve developing hybrid models that merge symbolic reasoning and neural representation and incorporating explainable AI platforms. The author also discusses the need for stakeholder engagement in co-designing these systems to achieve better usability and accountability in legal environments.

Study in [7] explores the deployment of an LLM-driven chatbot in higher education to assist with databases and information systems. The system uses GPT-based interaction to offer academic help to students. The research sheds light on challenges such as maintaining factual accuracy, minimizing hallucinations, and capturing user intent concerns equally relevant to legal NLP systems. The study suggests improvements through continuous retraining based on user interactions and proposes user feedback mechanisms to align chatbot responses with pedagogical goals. Its findings are transferable to legal applications where educational or support chatbots are in use.

Paper [8] introduces a novel multi-agent reinforcement learning (MARL) approach to enhance retrieval-augmented generation. It presents a structured agent framework where each agent specializes in a subtask (e.g., document retrieval, summary generation, and feedback aggregation), improving performance on multi-hop reasoning tasks. The authors evaluate their model on benchmarks like HotpotQA and propose applications in complex domains such as legal research, where reasoning over long, interconnected documents is crucial. Identified limitations include model scalability and coordination efficiency, which the authors suggest can be addressed by hierarchical agent modeling and inter-agent reward balancing.

Study in [9] presents a high-level overview of the Retrieval-Augmented Generation model and its relevance in domains requiring up-to-date and context-specific knowledge, such as law. The

authors emphasize that static LLMs lack temporal awareness and context adaptability, which RAG can potentially overcome. Key gaps discussed include the dependency on retrieval quality, difficulty in aligning retrieved and generated content, and absence of user-specific personalization. Proposed improvements involve query re-ranking, memory-augmented RAG pipelines, and feedback-driven retrieval enhancement. The paper encourages further empirical validation in high-risk domains, especially legal, healthcare, and finance.

"LawPal" in [10] introduces a RAG-based legal assistant system tailored to the Indian legal landscape. The system is designed to offer simplified legal advice, document summaries, and statute lookups. It integrates a retriever module based on BM25 and a generator based on GPT-3.5. The paper discusses the challenges of dealing with multilingual and multi-jurisdictional legal documents in India. Key limitations include limited context handling, slow performance, and lack of continuous feedback adaptation. The authors suggest incorporating court database APIs and enhancing multilingual support with regional language LLMs.

Paper [11] divides legal NLP applications into Legal Search, Legal Document Review, and Legal Prediction, with some of its applications being document retrieval, contract analysis, summarization, and prediction of case outcome. The authors note a notable expansion in quantity of studies, research sophistication, and multilingualism. Findings demonstrate the robust performance of domain-specific fine-tuned LLMs such as LEGAL-BERT in outperforming general models on different tasks. Major challenges involve data sparsity, interpretability, and tailoring models to intricate legal scenarios. The research also identifies growing trends and research needs, providing directions for enhancing AI-driven legal practice.

Paper [12] proposes multi-document handling via LLMs. It encourages comparative studies with GraphRAG to assess latency and performance. Though financial in nature, its techniques are relevant to legal domains dealing with contracts or regulations. The LAWSUIT dataset [13] provides expert-written summaries of Italian constitutional court verdicts. Gaps identified include linguistic biases, lack of graph-based summarization, and the need for corpus-level knowledge extraction. It encourages generalizing to low-resource languages.

InLegalLLaMA, [14] presents a knowledge-augmented large language model specifically designed for Indian Legal Text Analytics (LTA) tasks. The model utilizes a domain-knowledge legal graph constructed from Indian court judgments and statutes by incorporating structured triples into LLaMA-2 through ongoing pretraining and instruction tuning. It shows enhanced performance compared to baseline LLaMA-2 and Flan-T5 in triple prediction and rhetorical role

classification tasks, thereby showing the advantage of domain adaptation. The system overcomes issues such as multilingual content, difficult legal language, and unstructured information in Indian legal documents. Results show promise for broader applicability in applications such as case similarity detection, judgment summarization, and legal QA in the Indian legal ecosystem.

Paper [15] suggests a RAG framework for improving LLMs to respond to Canadian criminal law questions with an attempt to reduce hallucination problems. The architecture transcribes the Criminal Code to text, vectorizes it through representations, stores it in a Milvus vector database, and retrieves chunks that are relevant to LLM-generated responses. Gaps that were identified are the lack of comprehension of complicated or indirect questions, small size of the evaluation dataset, and lack of a standardized legal benchmark for RAG evaluation. The research also identifies usability constraints based on the absence of a graphical interface and only one legal domain coverage. It is recommended for future work to improve legal coverage, develop standardized test datasets, enhance context retrieval accuracy, and release the model on user-friendly platforms for legal experts and the general public.

Study in [16] introduces a thorough investigation on how search and retrieval strategies can be maximized for RAG systems with an emphasis on retrieval accuracy, efficiency, and relevance balance. The paper analyses different retriever architectures, ranking algorithms, and indexing techniques and puts forward a hybrid retrieval model that blends sparse and dense retrieval to provide better performance. The key gaps recognized are the absence of retrieval optimization by domain, restricted management of noisy or imperfect queries, and no adaptive retrieval mechanisms that adapt to query difficulty. Future research includes the development of adaptive hybrid retrievers, combining retrieval personalization, and the construction of benchmark datasets corresponding to real-world RAG application settings to evaluate them more thoroughly.

The systematic review of XAI [17] in law identifies the inadequacy of generic models across legal subdomains. It calls for crafting tailored explanations and segment-specific explainability to meet varying user needs in legal decision-making. Paper [18] discusses optimization of search and retrieval in RAG systems, highlighting major gaps in current work. The paper sheds light on under researched conditions on which Long Context(LC) or RAG operate optimally and proposes that an integration of the two could unlock their respective strengths. Discoveries on chunking techniques show optimal breakdown of extended contexts is unresolved, so more studies are needed on its effect on the accuracy of retrieval. For future research, the authors suggest the

development of integrated LC–RAG hybrid models to optimize performance on intricate question answering tasks.

The research [19] addresses fine-tuning the Mistral model for legal summarization but does not examine the impact of fine-tuning other LLMs, including GPT-4, perhaps indicating a lack of insight into model-specific variations in performance in the legal domain. It also points to the recurrent problem of hallucination in summaries produced, especially in the fine-tuned Mistral, which compromises factuality. The authors suggest future work to study fine tuning on multiple models of different types and to design strategies like enhancing training data quality or optimizing architectures to reduce hallucination errors and increase the reliability of legal summarization systems.

This comprehensive [20] survey identifies summarization challenges such as fluency, domain alignment, and deep learning integration. The paper calls for readable and legally accurate summaries through hybrid techniques and unsupervised models. Study in [21] proposes an NLP-powered legal query system aimed at improving accessibility to legal information. While practical in nature, it lacks an in-depth evaluation of explainability, multilingual capabilities, and long-document handling in its framework. Study in [22] explains various RAG approaches when integrated with LLMs. The authors identify issues like factual grounding, query disambiguation, and coherence. They suggest improving hybrid retrieval pipelines and enhancing long-context synthesis.

Paper [23] offers a comprehensive overview of RAG for LLMs, covering architectures, benchmarks, and application areas. Gaps include limited analysis of domain-specific datasets and insufficient attention to interpretability in legal applications. This paper [24] proposes ensemble learning for legal document similarity matching. While promising in matching efficiency, it lacks integration with generation-based systems like RAG and needs further testing on multilingual and real-world legal datasets.

Paper [25] presents LawBot, a tool based on AI developed to help users comprehend real estate legal documents, measured in terms of precision, recall, and F1 scores. Though efficient in enhancing legal understanding, the research identifies aspects of improvement such as more user-friendly user interaction design, inclusion of real-time legal updates, and increased evaluation measures that track contextual and legal accuracy. The technology has potential for other applications in addition to real estate, with versatility in applying to fields including family, corporate, and intellectual property law. Future research is proposed to involve user feedback

mechanisms for iterative refinement and to expand the tool's coverage across legal fields for higher utility and impact.

This research [26] canvasses the position of XAI in legal scenarios with a focus on the indispensable requirement of transparency in Deep Neural Network (DNN) models to facilitate accountability and trust among legal experts. It points to a gap in reconciling intricate AI models to be more interpretable to stakeholders in the legal field and emphasizes the need for methodologies that are capable of dealing with complex and dynamic legal domains where rules can be conflicting or continually changing. Future studies are suggested to establish new techniques of analysis and representation specifically for such sophisticated legal cases in order to further increase the applicability of AI in practical applications for the legal field.

Paper [27] proposes BM25 and BERT-based clustering for summarizing legal documents. Future directions include enhancing readability and improving cluster combination methods to generate more coherent summaries. The paper [28] introduces an LLM-based legal aid for governance activities, proving the system's ability to provide plausible responses but reporting occasional mistakes and vagueness, particularly with intricate legal questions. Future research focuses on improving evaluation mechanisms to enhance the accuracy, reliability, and clarity of responses generated in governance-oriented legal settings.

Paper [29] reviews current legal summarization methods. Gaps include lack of hybrid extractive-abstractive models, minimal use of unsupervised learning, and domain-specific model requirements. Knowledge representation and evaluation metrics are also proposed for future work. Work in [30] explores RAG survey discusses ambiguity handling, ethical issues, and coherence challenges. It recommends improved context integration, scalable solutions, and attention to societal impact of AI in law. Paper [31] reviews ethical perspectives of XAI. Gaps include user perspective understanding and industrial implementation. The study advocates for multimodal integration, privacy techniques, and fair AI deployment.

Paper [32] identifies the need for continuously updated explanations in XAI. It proposes interactive systems, hybrid explanation models, and adaptive context-driven explanations for robust legal AI applications. The research [33] examines retrieval-based evaluation approaches for LLMs in the case of Korean legal question answering, and it cites hallucination as a significant challenge that erodes factual accuracy. Future research is suggested for working towards reducing such inaccuracies so that responses generated by them are not only coherent but factually accurate too in legal contexts.

Paper [34] discusses challenges in summarizing unstructured, native-language legal texts. It suggests incorporating argumentative role classification and improving legal domain-specific summarization frameworks. Paper [35] presents the Indian Legal Corpus (ILC), a large dataset intended for summarizing Indian court proceedings, comprising court case files accompanied by manually prepared summaries. The dataset covers the lack of domain-specific resources for legal NLP in India and facilitates research on both extractive and abstractive summarization. Gaps found are scarce coverage of various case types and the requirement for better annotation consistency. Directions for future research include augmenting the corpus to encompass wider legal areas, improving annotation quality, and using the dataset to support advanced tasks like legal question answering and argument mining.

Paper [36] emphasizes the value of SHAP explanations but suggests exploring other methods. Future research should evaluate how enjoyment and satisfaction (hedonic factors) influence user trust in XAI. Paper [37] recommends contrastive learning and further neural architecture studies to improve summarization fidelity and relevance for Greek case law. Paper [38] review highlights the need for a comprehensive taxonomy in XAI. It also calls for integration with deep learning systems and systematic feedback loops to enhance trust and utility. Work in [39] proposes a high-precision two-stage summarization framework. Issues include handling long texts, redundancy control, and keyword integration. Coverage mechanisms need refinement for improved abstraction.

Paper [40] explores multilingual summarization via transfer learning. Future research should involve LLM experiments like ChatGPT and explore token-level embeddings for improved legal adaptation. Work in [41] introduces LEVEN, a massive Chinese legal event detection dataset, which includes annotated legal documents with event types, arguments, and triggers to support information extraction in the legal context. The paper capitalizes on BERT-based models and other neural architectures for event detection tasks and exhibits superior performance compared to standard baselines. Gaps that have been identified are difficulties in processing complicated event structures, vague triggers, and cross-document event reasoning. Future work directions include creating stronger models for in-depth legal event understanding, increasing the dataset to span a variety of case types, and investigating multilingual and cross-jurisdictional event detection for more universal use cases.

Paper [42] addresses the growing interest in prompt engineering for legal AI applications, especially in multilingual contexts. The authors analyse how current models handle legal

judgement prediction and observe a critical gap: completions generated by language models often lack informative richness. Specifically, they highlight that most outputs fail to explicitly list violated legal articles or provide explanatory justifications for the predictions made. This indicates a fundamental shortfall in current prompt-based methods, which tend to focus on outcome generation rather than legal reasoning. The study implies that for multilingual and multi-jurisdictional systems to be effective, they must move beyond surface-level predictions and incorporate mechanisms that generate detailed, reasoned, and article-specific outputs. The authors suggest that future models should integrate deeper legal ontologies and structured templates that guide the generation of more interpretable and legally-grounded completions, thus improving transparency and utility for practitioners.

Work in [43] evaluates the existing domain-independent models, such as BART and PEGASUS, for summarizing Indian legal documents. They observe that these models, though state-of-the-art in general NLP, exhibit poor performance when applied to the highly formalized and structurally unique nature of legal texts. One notable issue raised is the inadequacy of traditional evaluation metrics, like ROUGE, in capturing the effectiveness of legal summarization, since these metrics fail to reflect the semantic and argumentative value essential in legal summaries. The paper proposes the need for tailored evaluation frameworks and benchmarks specifically crafted for the legal domain. In terms of future directions, the authors identify multiple areas of research: enhancing model adaptation, including fine-tuning and domain-specific pretraining, to better suit the stylistic and terminological features of legal texts, developing hybrid models that combine extractive and abstractive summarization approaches; and incorporating legal knowledge bases and ontologies to improve contextual understanding

Paper [44] proposes an evaluation framework for legal document summarization, emphasizing that standard metrics like ROUGE and BLEU fail to capture the intent of legal summaries. It highlights the need for intent-based metrics, suggests cross-jurisdictional application, and recommends integrating advanced models like Legal-LED and LetSum to enhance summarization effectiveness and domain adaptability. Paper [45] presents Pile of Law, a massive, morally filtered legal text dataset aimed to facilitate the development of responsible AI for the legal profession. The research prioritizes careful filtering of data to reduce risks to privacy, bias, and abuse, while providing high-quality, varied legal content. Gaps identified are the requirement of enhanced bias detection, extended jurisdictional scope, and more sophisticated filtering mechanisms. Future efforts seek to increase dataset scale, improve ethical protections, and create standards for measuring responsible AI in legal NLP applications.

The Paper [46] introduces deep learning methods for legal text summarization and defines major gaps such as the scarcity of labeled data, dependence on generic methods, and lack of application of domain-specific adaptations. It recommends combining topic modeling and latent semantic analysis with neural networks to enhance semantic understanding. The research also emphasizes the necessity for legal-specific evaluation metrics and greater investigation of abstractive or hybrid summarization methods adapted to legal domains.

Paper [47] surveys deep learning methods for legal text summarization and identifies challenges like going to maintain coherence, minimizing redundancy, and the computational expense of abstractive types. It demands hybrid strategies that find a balance between quality and efficiency. The research emphasizes the need for XAI, customized summarization, and the use of semantic and discourse-level techniques. New transformer-based architectures should be developed by future work to enhance summarization results in the legal area as well.

Work in [48] discusses deep learning-based methods for legal text summarization in order to overcome issues of being coherent in multi-document summaries while minimizing redundancy. It observes that abstractive models such as pointer-generator networks are computationally costly and posits the necessity of efficient hybrid extractive–abstractive models. The research puts high value on the explainability of deep learning models to establish trust in generated summaries and looks into the possibility of user-centric summarization for particular needs. It also underlines the contribution of semantic, pragmatic, and discourse-level analysis to enhancing summarization quality and relevance. Future research is suggested to explore novel deep learning architectures, especially more sophisticated transformer models, and incorporate semantic analysis in order to generate better-quality, context-sensitive legal summaries.

Paper [49] summarizes information extraction (IE) methods for the analysis of legal documents highlighting KBP system improvement as necessary for unstructured legal texts. It underscores difficulties in constructing tagged corpora for deep learning and proposing to explore semi-supervised approaches to minimize manual effort. The research also identifies the possibility of open information extraction (OIE) to process legal text more adaptively. Future research focuses on combining predictive analytics to predict legal outcomes and improve decision-making instruments in legal practice.

Research in [50] investigates long document summarization in low-resource domains with pretrained models such as BART and GPT-2. It mentions that BART works poorly with lengthy texts, which indicates a require improved model adjustment or novel architectures. The authors

also suggest a GPT-2-based approach to salience detection, which leverages identification of important sentences but yet remains subject to improvement. They also point out that using ROUGE-L alone might not accurately reflect summary quality, particularly for legal content, and suggest considering richer evaluation measures in the future.

Paper [51] offers an informative critique of current summarization methods in the legal area and presents DELSumm, an unsupervised extractive summarization technique specific to Indian Supreme Court case files. One of the key contributions of the research is its focus on the integration of domain knowledge into summarization programs. DELSumm fills this void through giving preference to such portions, thus bringing its output in line with the expectation of legal professionals. This indicates the promise of unsupervised or hybrid methods for domains where labeled data is scarce. In addition, the authors propose that future research on: investigating abstractive summarization techniques able to synthesize and paraphrase across segments, increasing the system's generalizability across jurisdictions and legal systems and creating hybrid learning models that capitalize on the strengths of supervised and unsupervised methods be done. As a whole, this paper provides a solid ground for expert-matched, high-utility summarization in the legal field.

Paper [52] reviews progress in automatic text summarization (ATS) and highlights key challenges. It emphasizes the need for better abstraction techniques to create more insightful summaries and calls for improved evaluation metrics beyond standard ones like ROUGE. The paper also suggests integrating modern NLP and deep learning technologies to handle complex text structures. Future research should focus on new algorithms, enhanced feature extraction, and innovative approaches to advance summarization systems.

Work in [53] mainly focuses on extractive summarization, which involves selecting important sentences from the original document. However, they point out that abstractive summarization creating new sentences to better explain the content is much more challenging and requires deeper understanding and more computing power. They also highlight that current methods do not consider user preferences, which could make summaries more useful if customized. For future improvements, the authors suggest adding abstractive summarization techniques, making the process faster, tailoring summaries to users' needs, making the system explainable, using advanced NLP models like transformers, and adapting the methods specifically for legal language. These steps could help create clearer, more relevant, and trustworthy summaries for legal documents.

Paper [54] focused on certain types of legal documents and found that making AI more explainable helps users trust the system and complete tasks faster. However, they suggest future research should test explainability across different legal documents and contexts to see if it works well everywhere. They also emphasize designing explainability tools that fit the needs and thinking styles of legal professionals better. Additionally, the study recommends exploring how explainability affects the quality of summaries and accuracy of user judgments. Future work could also compare more explainability methods and look at how these features can be combined with other AI tools to improve legal workflows overall.

Gaps Identified from Literature Review:

• Better Integration of Extractive and Abstractive Summarization Necessary

Whereas extractive techniques guarantee factual accuracy and abstractive techniques offer smooth, human-sounding summaries, there is a common failing in most legal summarization models to bring the two together in an optimal manner. This, in turn, tends to produce either overly formal or fact-drifting outputs. Hybrid summarization techniques [2][29][30][51] have been studied to combine the powers of both to greatly enhance legal summary quality, but this avenue is yet to be fully explored in existing systems.

• Hallucination Issues

Even with recent improvements in Retrieval-Augmented Generation (RAG), hallucination that generated text contains unsubstantiated or made-up information is still an ongoing issue [33][15][19][49]. In law, these mistakes can have serious effects on trust and usability. Better grounding methods, fact-checking layers, and retrieval–generation alignment approaches should be employed to maintain outputs close to the original material.

• Redundancy in Extractive Summarization

Extractive summarization methods tend to replicate redundant content found in legal documents, resulting in inflated and less meaningful outputs [2]. Although redundancy reduction methods are available, their application is rare in legal NLP systems, which reflects the importance of specialized procedures that eliminate redundancy without sacrificing key legal sense.

• Multilingual and Cross-Border Legal Case Handling Limitations

Most existing models are cut out for single-language, single-jurisdiction cases and therefore do not apply to the increasing multilingual and cross-jurisdiction legal cases [35][2][40][11]. This shortcoming emphasizes the necessity of models capable of taking in multi-language inputs,

normalizing legal vocabulary across jurisdictions, and keeping contextual validity across varied legal systems.

• Insufficiency of Customized XAI Approaches to Legal Document Formulation

XAI techniques have the tendency not to capture the organized patterns of reasoning and hierarchical structure of legal texts [32]. Without explanations rooted in how lawyers' reason about statutes, precedents, and arguments, system outputs stand the risk of being less credible and more difficult to verify. Legal reasoning adapted to XAI may enhance interpretability and uptake.

• Lack of Adequate Keyword Embedding in Summarization

Keywords are instrumental in the retrieval and contextualization of legal information but are often not incorporated well into summaries [39]. This absence has the potential to result in summaries that exclude legally relevant terms, decreasing accuracy in case-specific or statute-centered situations.

• Restricted Long-Context Comprehension and Management of Long Documents

Legal briefs tend to be hundreds of pages long, with interdependent sentences and citations [10][35][39]. Modern models tend to lose context over long text spans, generating abstractions that omit major cross-references or high-level arguments. Better long-context modelling is needed to handle such complicated material.

Significance of Study

The legal field is dominated by complicated, long, and domain-specific texts that require accurate interpretation and summarization to support sound decision-making. Conventional methods are usually inadequate in describing the complex organization, subtle vocabulary, and context dependency present in legal texts [2][6][35]. This research aims to investigate and examine current legal summarization methods and corpora in order to determine strengths, weaknesses, and possibilities of innovation in this field.

Drawing on these findings, the study suggests designing a hybrid extractive–abstractive summarization model coupled with XAI techniques to improve factual accuracy, long-context coherence, and grounding of produced outputs on accessed content [7][10][45]. Coupling XAI guarantees that legal experts will be able to track summaries back to their sources, enhancing transparency, interpretability, and confidence in system-generated results [13][32][39].

To guarantee robustness and practical applicability, the system will be tested by combining quantitative evaluation measures with expert legal responses to ensure outputs are to professional standards yet still interpretable [6][7][35]. In addition, the framework is also designed to facilitate multilingual summarization of legal documents to meet emerging demands for cross-jurisdictional flexibility and greater accessibility [14][29][41][55].

In summary, this work not only contributes to the state-of-the-art in legal NLP, but it also provides a practical, transparent, and multilingual application for enhanced access to justice, efficient legal research, and assistance with decision-making in national settings.

5. National Status:

In India, the integration of Artificial Intelligence (AI) into the legal domain is still in its nascent but rapidly expanding phase. A number of academic and industry-led initiatives have emerged in areas such as legal text analytics, judgment retrieval, and basic summarization. Premier institutions like the Indian Institutes of Technology (IITs), National Law Universities (NLUs), and Indian Institutes of Information Technology (IIITs) have begun exploring Natural Language Processing (NLP) applications in the legal field, including legal document classification, entity recognition, and extractive summarization.

However, advanced frameworks that incorporate hybrid summarization (extractive + abstractive), legal keyword-guided processing, and domain-specific question answering remain relatively underexplored in the Indian context. Most existing tools developed by Indian legal tech startups (e.g., CaseMine, NearLaw, LegalMind) still rely heavily on rule-based methods or surface-level extractive summarization. The use of XAI for enhancing transparency in legal outputs and multilingual summarization for addressing cross-jurisdictional legal documents is even more limited, indicating a significant opportunity for deeper academic and industry collaboration.

Notably, initiatives like SUPACE (Supreme Court Portal for Assistance in Court Efficiency), launched by the Supreme Court of India, demonstrate growing institutional interest in leveraging AI to improve judicial processes. However, the adoption of deep learning-based legal summarization systems with integrated explainability, domain-aware QA, and multilingual capabilities remains minimal highlighting the need for research frameworks like the one proposed in this study.

6. International Status:

Globally, the application of Artificial Intelligence (AI) in the legal domain has gained significant traction, particularly in countries such as the United States, the United Kingdom, European Union member states, and China. There has been considerable progress in legal Natural Language Processing (NLP), including abstractive summarization, hybrid summarization, legal-specific large language models, and explainable AI (XAI).

Academic Research: Leading institutions like Stanford University, Harvard, University College London, and Tsinghua University are conducting advanced research in legal document understanding, domain-specific question answering, and intelligent legal summarization. Current advancements in legal technology are increasingly geared toward enhancing the accuracy, interpretability, and scalability of systems designed to process extensive legal documentation.

Industry Adoption and Innovations: Legal technology firms and startups worldwide are embracing advanced transformer-based models, combining both extractive and abstractive summarization techniques to improve outcome quality. Many of these solutions incorporate customized versions of large language models such as legal-specific adaptations of GPT-4, T5, and BERT for applications like document summarization, contract analysis, and regulatory compliance verification. In addition, several platforms integrate keyword-focused strategies and legal entity recognition to improve contextual relevance and precision.

Focus on Transparency and Responsible AI: With rising demands for responsible and trustworthy AI in the legal domain, international frameworks such as the EU AI Act and the OECD AI Principles have spurred increased emphasis on explainable XAI. These guidelines promote the development of legal tools that offer clear, justifiable outputs, addressing ethical concerns and fostering greater accountability. Tools such as LIME, SHAP, and emerging frameworks like LegalXAI are being customized to ensure transparency in legal reasoning and AI-generated outputs, fostering user trust in sensitive domains.

Multilingual Legal AI: The multilingual nature of legal systems within the European Union has catalysed the development of cross-lingual legal NLP tools and datasets. Projects such as MULAI (Multilingual Legal AI) and datasets like JRC-Acquis and EURO-LEX have facilitated significant advancements in multilingual summarization and legal translation, allowing AI systems to operate across diverse legal jurisdictions and languages.

7. Study Hypotheses:

In order to prove the efficiency of the suggested hybrid legal summarization system and determine whether it indeed provides quantifiable improvements compared to current legal NLP practices, the following hypotheses have been developed.

H_0 (Null Hypothesis):

The suggested hybrid framework does not demonstrate any significant improvement regarding factual accuracy, semantic coherence, reduction of hallucinations, transparency, or quality of multilingual summarization when compared to existing legal summarization system.

H_1 (Alternative Hypothesis):

The suggested hybrid framework results in distinct and quantifiable improvements in factual accuracy, semantic coherence, reduction of hallucinations, transparency, and multilingual summarization quality over existing legal summarization system.

This test of the hypothesis will establish whether the framework developed actually confers measurable benefits in legally inspired real-world applications. Controlled experiments will be carried out in which both the system proposed and baseline models are used under the same circumstances, with the same legal data sets. Representative performance metrics including ROUGE, BERTScore, factual consistency, hallucination rate, and expert-based interpretability measures will be compared.

Statistical testing procedures (e.g., paired t-test, Wilcoxon signed-rank test, or ANOVA) will be used to test results. In case improvements derived by the proposed framework are statistically significant, the null hypothesis (H_0) will be rejected in favor of the alternative hypothesis (H_1). This would validate that the proposed framework is effective and consistent with the outlined research goals.

8. Study Rationale:

Legal documents, often long and syntactically sophisticated, also differ across jurisdictions, posing a very tough analysis or summarization task through the traditional NLP approaches. As the legal field embraces AI tools ever more so for searching and deciding, the deficiencies in current summarization methods begin to emerge. Legal texts differ from other documents in demanding

factual accuracy and contextual pertinence while also demanding that the summaries they generate allow for clarity and explainability, criteria that many techniques currently fail to meet.

While the majority of extractive summarization methods merely consider surface-level content, they are incapable of comprehending the deeper semantics of legal text. Abstractive summarization techniques often ill-adapted to legal domain tasks may lead to faulty or even entirely erroneous interpretations, which is an all-but-absurd possibility for the legal domain. Admittedly, most generic language models constitute a less than perfect fit when it comes to deciphering the intricacies of legal terminology, syntax, and modes of reasoning. Furthermore, the opaqueness of numerous AI systems exacerbates these challenges, as they frequently fail to provide interpretable justifications for their generated outputs a vital need in contexts involving legal sensitivity. This complexity is heightened in multilingual or international environments, where vocabulary and legal frameworks differ significantly across jurisdictions.

To overcome these challenges, a multi-layer framework is proposed with systems aligned to the special needs of legal NLP. The hybrid summarization system works by interrelating extractive and abstractive summarization techniques to keep the summaries legally structured and yet essentials conveyed to the user. Keyword extraction enhancements make the summarization focus on legally significant terms that are important conceptually, thereby increasing the accuracy of summarization and also enhancing the question answering module on domain specifics.

The system enjoys legal enhancement in virtue of the legal domain language models, which strengthens its abilities to understand legal text and diminishes the threat of misinterpretation. Explainable AI mechanisms come with the system, making it possible to explain and verify every output that this system produces and enabling tracing of the rationale for every conclusion made by a legal expert. And due to multilingual capabilities built in, it could well be used in different legal jurisdictions, allowing for cross-border legal analysis and documentation in various languages.

As the merging of accuracy, clarity, legal reasoning, and interpretability, this framework stands as a giant leap for legal AI. It aims to support legal persons, researchers, and courts with accurate summaries and answers-with jurisdictional awareness-to foster the efficiency and transparency of the legal workflow increasingly subjected to digitization.

9. Research Methodology:

The methodology adopted in this project is designed to systematically process legal documents and generate high-quality summaries that are legally faithful, concise, and interpretable. Given the complex and verbose nature of legal language, the pipeline incorporates both NLP techniques and

domain-specific retrieval strategies to ensure factual consistency and legal accuracy. This section outlines the end-to-end workflow (Figure 1.3), beginning with input document pre-processing and progressing through entity recognition, summarization, explainability, and multilingual support.

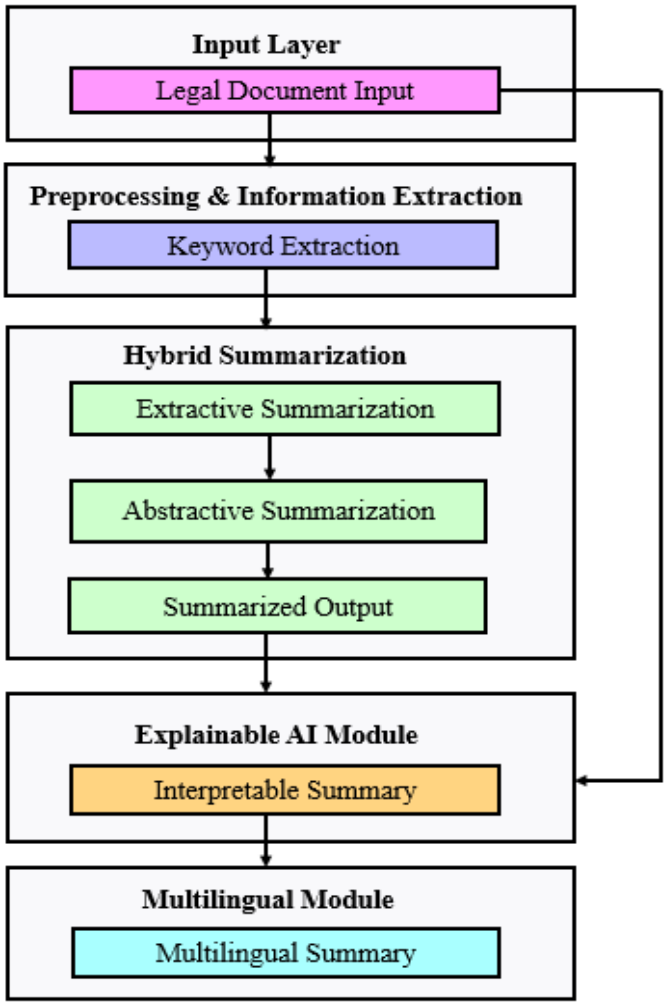

Fig 1.3: Proposed Methodology of Legal Document Summarization

The following components illustrate the sequential flow of the framework:

1. **Input Document**

The system input is a legal document, including contracts, court rulings, case overviews, terms and conditions, or legal notices, and an optional user query for question answering. The documents usually involve lengthy, specialist domain language with complicated sentence structures. The objective is to make them more concise and summarise them without losing their legal validity. The document is consumed in text or PDF form and serves as the input for stages in the pipeline that follow.

2. Pre-processing and Legal Keyword Extraction

Once the input document is received, the first processing step involves Pre-processing and Legal Keyword Extraction. This step leverages NLP models trained to detect domain-specific entities such as names of laws, clauses, organizations, individuals (plaintiff, defendant), dates, and locations. Additionally, a legal dictionary or rule-based extractor is employed to identify and extract important legal terms and keywords like "indemnity," "jurisdiction," or "breach of contract." This stage helps in understanding the core legal structure of the document, which is crucial for summarization and evaluation in later stages.

3. Hybrid Summarization

This module provides a combination of extractive and abstractive summarizations.

Extractive Summary: The extractive summarizer determines and collects the most germane sentences or phrases from the source content to produce a short summary. Unlike abstractive systems which create new text, the extractive method selects sentences directly from the source material. Such an approach ensures that no fictitious information is introduced, with all statements retaining their original legal phrasing. This mechanism serves two major purposes-i.e., (1) to provide a basic summary to contrast with the abstractive one; and (2) to establish evaluation criteria such as ROUGE, which analyses lexical overlaps between the two summaries.

Abstractive Summary: Next, the document is submitted for abstractive summarization, making use of state-of-the-art language generation models like BART or T5. The abstractive summarizer, unlike an extractive summarizer, generates novel sentences using linguistic patterns that the system has learned, allowing the system to paraphrase, condense, and rewrite long passages into shorter, clearer messages. This is especially beneficial in the legal field, where verbosity and duplicates frequently occur. The produced summary seeks to preserve the legal significance and context while enhancing comprehensibility for non-experts, including clients and junior legal assistants.

This hybrid summarizing module output integrates extractive and abstractive processes and is called a grounded summary. During this stage, the necessity for the correct legal format, the appropriate legal context, and freedom from error in the generated summary are underscored. The summary must have only the verifiable information obtained from the genuine legal document source, in order to maintain its legal integrity; readability is hence enhanced. This grounded output thereby moves on for evaluation, elucidation, and possible multilingual translation.

4. Explainable AI Module

The project incorporates an XAI module to emphasize the importance of transparency in legal contexts. This module identifies the words, phrases, or sections of the input document that have

the greatest influence on the final summary output. The model's decision-making process is revealed through visual explanations generated by the XAI module, enabling legal experts to verify and validate the summary and thereby increasing confidence in the system's output.

5. Validation Module

Following the XAI analysis, the summary undergoes a validation stage. This includes quantitative metrics such as ROUGE and BLEU to assess lexical overlap and fluency, and legal-specific measures to verify factual and legal accuracy. Expert feedback is also collected to ensure that the system's outputs are both correct and trustworthy. This step ensures the final multilingual summary is both accurate and transparent.

6. Multilingual Translation Module

To serve users across different regions of India, an additional multilingual translation module is included. The final summary is translated into multiple Indian languages using pre-trained models. This is particularly useful in multi-jurisdictional legal contexts within India, where clients and stakeholders may not be proficient in the original language of the legal document. The translation is executed meticulously to guarantee the preservation of legal meanings and terminologies across languages.

10. Statistical plan

- The statistical plan for this study involves a comparative evaluation of the proposed hybrid summarization framework against baseline models including standalone extractive summarization, and abstractive summarization systems.
- Legal datasets in multiple Indian languages will be used, to ensure robust and representative analysis.
- Performance evaluation will be conducted using established summarization metrics such as ROUGE (Recall-Oriented Understudy for Gisting Evaluation), BLEU (Bilingual Evaluation Understudy), and BERTScore, assessing content quality, fluency, and semantic similarity.
- Expert reviewers will provide qualitative feedback on the clarity and effectiveness of the XAI module in conveying the reasoning behind the system’s outputs.
- Statistical tests such as paired t-tests, Wilcoxon signed-rank tests, and ANOVA will be applied to determine the significance of performance differences across different model configurations and modules.

- The evaluation process will be replicated for legal documents in various Indian languages, allowing assessment of the framework's adaptability and effectiveness across multilingual legal contexts.

11. Limitations of the Study

Although the proposed framework is designed to advance the state of legal document summarization through hybrid techniques and explainable AI integration, certain constraints may affect its applicability and performance. The following limitations highlight potential challenges in data availability, model adaptability, and evaluation methodologies that could influence the generalizability and effectiveness of the system across diverse legal contexts.

- Limited Availability of Legal Datasets: The study may rely on a limited set of publicly available legal corpora, which may not fully represent the diversity of legal systems, jurisdictions, and languages. This can affect the generalizability of the model across different legal domains.
- Complexity of Long Legal Documents: Despite using long-context models, current Transformer-based architectures (even with RAG) have inherent limitations in processing extremely long documents without losing contextual relevance or coherence.
- Evaluation Challenges: Standard summarization metrics (e.g., ROUGE, BLEU) may not adequately capture legal correctness, factual grounding, or argument structure. While additional human evaluations are proposed, they are time-consuming and subjective, which may limit scalability.
- Explainability Trade-offs: Incorporating XAI often introduces a trade-off between interpretability and performance. Explaining complex decisions in a legally meaningful way remains a challenge, particularly when summarization involves multiple nested legal arguments.
- Multilingual Support Constraints: While the framework aims to support multilingual legal summarization, the effectiveness is limited by the availability and quality of pretrained domain-specific models in non-English languages.
- Jurisdictional Alignment Issues : Adapting models to handle jurisdiction-specific laws, terminologies, and interpretations requires fine-tuning on jurisdiction-specific data—which may not always be accessible or comprehensive.

12. Expected Outcomes of the Study

The outcome is the development and evaluation of a hybrid AI framework for legal document summarization that demonstrates the following key points:

- Improved Summarization Quality: The integrated hybrid summarization approach, which combines extractive and abstractive techniques, is anticipated to generate more coherent, semantically accurate, and legally meaningful summaries of long and complex legal documents. This method is expected to outperform standalone extractive or abstractive models in capturing both factual detail and legal context.
- Improved Explainability and Transparency: A XAI module is included in the proposed framework to offer understandable justifications for every summarization result. This feature is essential for building user confidence and ensuring compliance with ethical and legal standards, especially in cases where decisions carry significant legal or emotional weight.
- Multilingual Capabilities: The system adjusts for local legal language and contextual nuances, making it suitable for legal summarization in different Indian languages. This functionality supports diverse legal practices across India, making the framework applicable to multilingual legal settings within the country.
- Paving the Way for Future Legal AI Innovations: The design and insights from this study are set to guide future developments in legal AI systems. Possible applications include automating legal research, summarizing legal cases, conducting compliance assessments, and analysing legal policies. This can help create more efficient and accessible legal processes

13. Conclusion:

Trusted and clever summarization tools are required because of the steep rise in document amount and complexity within the legal industry. For their research, decision-making, and quick access to critical information, legal experts need these tools. This research offers an extensive system that uses keyword identification, and a hybrid summary technique, to generate brief, legally accurate, and contextually appropriate results. XAI is integrated into the plan to promote trust and responsibility through enabling users to comprehend how summaries are created, which is highly important within high stakes legal situations. The design also has a multilingual component that broadens its usability to various legal jurisdictions and provides greater accessibility for a diverse demographic of users. This architecture supports advanced AI solutions tailored for the legal sector by infusing precision, clarity, and linguistic flexibility. It manages to bridge law expertise and technical intelligibility, promoting the creation of an ethical, smart, and inclusive legal assistance framework.

14. Bibliography

[1] Tulay Turan et al “Summarization, Prediction, and Analysis of Turkish Constitutional Court Decisions with Explainable Artificial Intelligence and a Hybrid Natural Language Processing Method, Digital Object Identifier 10.1109/ACCESS. 2025

[2] Maha Ahmed Abdullah Albayati et al, “A Hybrid Transformer-Based Framework for Multi-Document Summarization of Turkish Legal Documents” ,2025, Digital Object Identifier 10.1109/ACCESS.2025.3545750

[3] Rahman S. M. Wahidur et al, “Legal Query RAG”, IEEE Access, 2025, Digital Object Identifier 10.1109/ACCESS.2025.3542125.

[4] Majjed Al-Qatf et al, “RAG4DS: Retrieval-Augmented Generation for Data Spaces -A Unified Lifecycle, Challenges, and Opportunities, IEEE Access, 2025, Digital Object Identifier 10.1109/ACCESS.2025.354538

[5] Marco Siino, “Exploring LLMs Applications in Law: A Literature Review on Current Legal NLP Approaches”, 2025, Digital Object Identifier 10.1109/ACCESS.2025.3533217

[6] Mahd Hindi, ”Enhancing the Precision and Interpretability of Retrieval-Augmented Generation (RAG) in Legal Technology: A Survey”, IEEE Access, 2025, Digital Object Identifier 10.1109/ACCESS.2025.3550145.

[7] Alexander Tobias Neumann et al, “An LLM-Driven Chatbot in Higher Education for Databases and Information Systems”, IEEE TRANSACTIONS ON EDUCATION, VOL. 68, NO. 1, FEBRUARY 2025.

[8] Y. Chen, L. Yan, W. Sun, X. Ma, Y. Zhang, S. Wang, D. Yin, Y. Yang, and J. Mao, "Improving Retrieval-Augmented Generation through Multi-Agent Reinforcement Learning," arXiv:2501.15228v1 [cs.CL], Jan. 2025. [Online]. Available: https://arxiv.org/abs/2501.15228

[9] A. Mishra and A. Gupta, “Retrieval Augmented Generation (RAG) Model, International Journal of Research Publication and Reviews, Vol 6, no 1, pp 4690-4693 January 2025

[10] D. Panchal, A. Gole, V. Narute, and R. Joshi, "LawPal: A Retrieval Augmented Generation Based System for Enhanced Legal Accessibility in India," arXiv:2502.16573v1 [cs.AI], Feb. 2025. Available: https://arxiv.org/abs/2502.16573v1

[11] Marco Siino et al, “Exploring LLMs Applications in Law: A Literature Review on Current Legal NLP Approaches”,2025,IEEE Access, Digital Object Identifier 10.1109/ACCESS.2025.3533217

[12] Shalin Shah, ”Multi-Document Financial Question Answering using LLMs”, arXiv:2411.07264v1 [cs.IR] 8 Nov 2024

[13] Luca Ragazzi et al, “LAWSUIT: a LArge expert-Written SUmmarization dataset of ITalian constitutional court verdicts”, Springer, 2024, https://doi.org/10.1007/s10506-024-09414-w.

[14]Sudipto Ghosh, “InLegalLLaMA: Indian Legal Knowledge Enhanced Large Language Model”, The First international OpenKG Workshop: Large Knowledge-Enhanced Models, August 03, 2024, Jeju Island, South Korea.

[15] Ruiteng Li, “Enhanced Large Language Models-based Legal Query Responses through Retrieval Augmented Generation”, Proceedings of the 2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024), https://doi.org/10.2991/978-94-6463-512-6_27.

[16] Alexandria Leto, “Toward Optimal Search and Retrieval for RAG”, a rXiv:2411.07396v1 [cs.CL] 11 Nov 2024.

[17] G. Joshi, S. Mali, and P. G. Fegade, "A Systematic Review on Explainable AI in Legal Domain," 2024, International Journal for Research in Applied Science & Engineering Technology ,DOI: https://doi.org/10.22214/ijraset.2024.61736.

[18] Xinze Li, “Long Context vs. RAG for LLMs: An Evaluation and Revisits”, arXiv:2501.01880v1 [cs.CL] 27 Dec 2024.

[19] Mourad Heddaya, “CASESUMM: A Large-Scale Dataset for Long-Context Summarization from U.S. Supreme Court Opinions”, arXiv:2501.00097v1 [cs.CL] 30 Dec 2024.

[20] Rajendra A B et al, “A SURVEY ON SUMMARIZATION OF LEGAL TEXTS”, International Journal of Scientific Research in Engineering and Management (IJSREM), 2024. DOI: 10.55041/IJSREM29286

[21] Pushpa R N et al, “AI Powered Legal Querying System using NLP”, International Journal for Research in Applied Science &Engineering Technology (IJRAS), 2024.

[22] Muhammad Arslan et al, “A Survey on RAG with LLMs”, Elsevier, 2024, 10.1016/j.procs.2024.09.178

[23] Yunfan Gao et al, “Retrieval-Augmented Generation for Large Language Models: A Survey”, arXiv:2312.10997v5 [cs.CL] 27 Mar 2024.

[24] Aman Fan et al, “Legal Document Similarity Matching Based on Ensemble Learning”, IEEE Access, 2024, Digital Object Identifier 10.1109/ACCESS.2024.3371262

[25] Surbhi Dilip Pagar, “LawBot : From Documents to Answers, Unveiling A New Era In Real Estate Legal Assistance”, International Journal Of Scientific Research In Engineering And Management 08(05):1-5, 2024, DOI:10.55041/IJSREM34216.

[26] K. Mcgregor, S. M. Muddamsetty et al "Explainable AI and Law: An Evidential Survey," 2023. DOI: https://doi.org/10.1007/s44206-023-00081-z.

[27] L. Bonalume and K. Becker, "BB25HLegalSum: Leveraging BM25 and BERT-based clustering for the summarization of legal documents," 2023.: DOI: https://doi.org/10.26615/978-954-452-092-2_029.

[28] M. Mamalis and Evangelos, "A Large Language Model based legal assistant for governance applications,", 2023, DOI: https://arxiv.org/abs/AB0487824BFEBC3B809AD4F5BED52048.

[29] Sheetal Ajaykumar Takale, "A Survey of Legal Document Summarization Methods", International Journal of Advanced Research in Computer and Communication Engineering, 2023, DOI: 10.17148/IJARCCE.2023.1272.

[30] Shailja Gupta, Rajesh Ranjan, Surya Narayan Singh "A Comprehensive Survey of Retrieval-Augmented Generation (RAG): Evolution, Current Landscape and Future Directions"

[31] Heidi Vainio-Pekka, Mamia Ori-Otse Agbese et al "The Role of Explainable AI in the Research Field of AI, Ethics," 2023, ACM Transactions on Interactive Intelligent Systems, Vol. 13, No. 4, Article 26, https://doi.org/10.1145/3599974.

[32] Wenli Yang, Yuchen Wei, et al, "Survey on Explainable AI: From Approaches, Limitations and Applications Aspects", 2023, DOI: https://doi.org/10.1007/s44230-023-00038-y

[33] Cheol Ryu et al, "Retrieval-based Evaluation for LLMs: A Case Study in Korean Legal QA",

[34] Nikita, D. P. Rana and R. G. Mehta, "Research Challenges for Legal Document Summarization," IEEE World Conference on Applied Intelligence and Computing, 2023. doi: https://doi.org/10.1109/AIC.2023.50

[35] P. Trivedi, D. Jain, S. Gite, K. Kotecha, A. Bhatt, and N. Naik, "Indian Legal Corpus (ILC): A Dataset for A dataset summarizing Indian Legal Proceedings using Natural Language," 2023,https://doi.org/10.30919/es1022.

[36] P. Hamm, M. Klesel, P. Coberger, and H. F. Wittmann, "Explanation matters: An experimental study on explainable AI,", 2023, springer, https://doi.org/10.1007/s12525-023-00640-9.

[37] M. Koniaris, D. Galanis, E. Giannini, and P. Tsanakas, "Evaluation of Automatic Legal Text Summarization Techniques for Greek Case Law," 2023, DOI: https://doi.org/10.3390/info14040250.

[38] Saranya A.1, Subhashini R, "A systematic review of Explainable Artificial Intelligence models and applications: Recent developments and future trends",

[39] Yue Huang, Lijuan Sun, Chong Han and Jian Guo, "A High-Precision Two-Stage Legal Judgment Summarization", MDPI, 2023, https://doi.org/10.3390/math11061320.

[40] Gianluca Moro, "Multi-language transfer learning for low-resource legal case summarization", springer,2023, https://doi.org/10.1007/s10506-023-09373-81 3

[41] Feng Yao et al, “LEVEN: A Large-Scale Chinese Legal Event Detection Dataset”, arXiv:2203.08556v1 [cs.CL] 16 Mar 2022

[42] Dietrich Trautmann et al, “Legal Prompt Engineering for Multilingual Legal Judgement Prediction”, arXiv:2212.02199v1 [cs.CL] 5 Dec 2022

[43]Satyajit Ghosh et al, “Indian Legal Text Summarization: A Text Normalisation-based Approach”, rXiv:2206.06238v2 [cs.CL] 13 Sep 2022.

[44] Ankan Mullick et al, “An Evaluation Framework for Legal Document Summarization”, arXiv:2205.08478v1 [cs.CL] 17 May 2022.

[45] Peter Henderson, Mark S. Krass, Lucia Zheng, Neel Guha , “Pile of Law: Learning Responsible Data Filtering from the Law and a 256GB Open-Source Legal Dataset”, arXiv:2207.00220v2 [cs.CL] 29 Nov 2022.

[46] Deepa Anand, “Effective deep learning approaches for summarization of legal texts”, https://doi.org/10.1016/j.jksuci.2019.11.015

[47] Reshma Sheik et al, “Deep Learning Techniques for Legal Text Summarization”, ResearchGate, 2022, https://www.researchgate.net/publication/361444272

[48] Kalpesh Krishna et al, “Hurdles to Progress in Long-form Question Answering”, arXiv:2103.06332v2 [cs.CL] 19 May 2021.

[49] Ashwini V. Zadgaonkar, Avinash J. Agrawal, “An overview of information extraction techniques for legal document analysis and processing”, International Journal of Electrical and Computer Engineering (IJECE),2021, DOI: 10.11591/ijece.v11i6.pp5450-5457.

[50] Ahsaas Bajaj et al, “Long Document Summarization in a Low Resource Setting using Pretrained Language Models”, arXiv:2103.00751v1 [cs.CL] 1 Mar 2021.

[51] Paheli Bhattacharya et al, “Incorporating Domain Knowledge for Extractive Summarization of Legal Case Documents”, arXiv:2106.15876v1 [cs.CL] 30 Jun 2021

[52] M. F. MRIDHA , “A Survey of Automatic Text Summarization: Progress, Process and Challenges”, 2021, IEEEE Access, Digital Object Identifier 10.1109/ACCESS.2021.3129786.

[53] Rahul C Kore, Prachi Ray, Priyanka Lade, Amit Nerurkar, "Legal Document Summarization Using NLP and ML Techniques," *International Journal of Engineering and Computer Science*, vol. 9, no. 5, pp. 4488-4495, 2021. DOI: 10.18535/ijecs/v9i05.4488.

[54] Milda Norkute, et al, ”Towards Explainable AI: Assessing the Usefulness and Impact of Added Explainability Features in Legal Document Summarization”, 2021, Yokohama, Japan. ACM, New York, https://doi.org/10.1145/3411763.3443441.

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Monisha Linkesh
Assignment title:	Assignment1
Submission title:	Design Framework For Legal Document Summarization
File name:	Topic_Approval_Report-_Monisha_Linkesh.docx
File size:	391.41K
Page count:	32
Word count:	10,189
Character count:	65,979
Submission date:	17-Aug-2025 12:33PM (UTC+0530)
Submission ID:	2730359409

ABSTRACT

A significant rise in the volume and complexity of legal documents has created an urgent need for automated solutions to help legal professionals efficiently extract and summarize essential information. The legal domain is characterized by highly detailed, context-rich, and jurisdiction-specific language, where accurate interpretation and factual precision are crucial. Traditional summarization methods often fall short in addressing the unique challenges of legal texts, such as intricate contextual reasoning, Indian legal terminology, multilingual requirements, and the necessity for complete grounding in the source material.

This study begins by exploring and analysing existing legal summarization approaches and datasets, with a focus on their applicability to Indian legal content. The study builds on the limitations observed in prior work, including inadequate integration of extractive and abstractive methods, hallucination issues, redundancy in extractive techniques, and insufficient multilingual support. To address these challenges, the proposed framework introduces a hybrid summarization pipeline. It integrates keyword-driven legal information extraction with a hybrid module that combines extractive precision and abstractive fluency, while an Explainable AI (XAI) layer ensures interpretability by highlighting the reasoning behind the generated summaries.

A validation stage evaluates system performance using both quantitative metrics (ROUGE, BLEU, BERTScore) and expert feedback from legal practitioners, ensuring accuracy, transparency, and reliability. Finally, the framework extends to multilingual summarization across Indian languages, addressing the increasing demand for accessibility in diverse legal and linguistic contexts.

By combining hybrid summarization, domain-specific keyword embedding, explainable reasoning, and multilingual adaptability, the framework aims to improve accuracy, interpretability, and compliance in legal summarization. This contribution lays the groundwork for AI-powered tools to support legal research, multilingual case analysis, and professional practice in India, thereby addressing critical gaps in existing literature and enabling more trustworthy legal AI applications.

i

Copyright 2025 Turnitin. All rights reserved.

Monisha Linkesh

Design Framework For Legal Document Summarization

Assignment1

MultiCon25

Thakur College of Engineering and Technology (TCET)

Document Details

Submission ID
trn:oid:::1:3314833206

Submission Date
Aug 17, 2025, 12:31 PM GMT+5:30

Download Date
Aug 17, 2025, 12:36 PM GMT+5:30

File Name
Topic_Approval_Report-_Monisha_Linkesh.docx

File Size
391.4 KB

32 Pages

10,189 Words

65,979 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer
Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify writing that is likely AI generated as AI generated and AI paraphrased or likely AI generated and AI paraphrased writing as only AI generated) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

Design Framework For Legal Document Summarization

ORIGINALITY REPORT

5%	4%	4%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Maha Ahmed Abdullah Albayati, Kürşat Mustafa Karaoğlan, Oğuz Findik. "Towards efficient multi-legal document summarization: An ensemble approach for Turkish law", Engineering Science and Technology, an International Journal, 2025 Publication	1%
2	btu.edu.ge Internet Source	<1%
3	huggingface.co Internet Source	<1%
4	www.mdpi.com Internet Source	<1%
5	repository-tnmgrmu.ac.in Internet Source	<1%
6	ceur-ws.org Internet Source	<1%
7	Namita Singh Malik, Elizaveta A Gromova, Smita Gupta, Balamurugan Balusamy. "Legal	<1%

Analytics - The Future of Analytics in Law", CRC Press, 2022
Publication

8	arxiv.org Internet Source	<1%
9	riunet.upv.es Internet Source	<1%
10	Ritesh Maurya, Nikhil Rajput, M. G. Diviit, Satyajit Mahapatra, Manish Kumar Ojha. "Exploring the potential of lightweight large language models for AI-based mental health counselling task: a novel comparative study", Scientific Reports, 2025 Publication	<1%
11	pdffox.com Internet Source	<1%
12	www.greenpeace.org Internet Source	<1%
13	Candida M. Greco, Andrea Tagarelli. "Bringing order into the realm of Transformer-based language models for artificial intelligence and law", Artificial Intelligence and Law, 2023 Publication	<1%
14	discovery.researcher.life Internet Source	<1%

www.nortonrosefulbright.com

15	Internet Source	<1%
16	Aman Fan, Shaoxi Wang, Yanchuan Wang. "Legal Document Similarity Matching Based on Ensemble Learning", IEEE Access, 2024 Publication	<1%
17	www.ijcaonline.org Internet Source	<1%
18	1library.net Internet Source	<1%
19	docslib.org Internet Source	<1%
20	impressions.manipal.edu Internet Source	<1%
21	indianexpress.com Internet Source	<1%
22	Juli Bakagianni, Kanella Pouli, Maria Gavriilidou, John Pavlopoulos. "A systematic survey of natural language processing for the Greek language", Patterns, 2025 Publication	<1%
23	Marco Siino. "Exploring the use of LLMs in the Italian legal domain: A survey on recent applications", Computer Law & Security Review, 2025 Publication	<1%

24	Saloni Sharma, Piyush Pratap Singh. "Advancements in legal text summarization: integrating InLegalBERT for effective extractive summarization", International Journal of System Assurance Engineering and Management, 2025 Publication	<1%
25	oulurepo.oulu.fi Internet Source	<1%
26	stax.strath.ac.uk Internet Source	<1%
27	www.internationalpubls.com Internet Source	<1%
28	"Applied Cognitive Computing and Artificial Intelligence", Springer Science and Business Media LLC, 2025 Publication	<1%
29	"Explainable Artificial Intelligence", Springer Science and Business Media LLC, 2024 Publication	<1%
30	"Legal Knowledge and Information Systems", IOS Press, 2024 Publication	<1%
31	Masike Malatji. "Augmented Intelligence Framework for Human–Artificial Intelligence	<1%

Teaming in Cybersecurity", Human-Centric Intelligent Systems, 2025
Publication

32	docs.google.com Internet Source	<1%
33	imi.hwg-lu.de Internet Source	<1%
34	iris.polito.it Internet Source	<1%
35	lume.ufrgs.br Internet Source	<1%
36	www.frontiersin.org Internet Source	<1%
37	www.grafiati.com Internet Source	<1%
38	www.ijisrt.com Internet Source	<1%
39	www.researchgate.net Internet Source	<1%

Exclude quotes	On	Exclude matches	< 8 words
Exclude bibliography	On		

University of Mumbai

Mahatma Jotirao Phule Bhavan, Vidyanagri, Santacruz (E), Mumbai - 400 098.

RECEIPT NO.: 3582 **DATE :** 22/08/2025

GENERAL FUND - I

Received from LINKESH MONISHA the sum of Rupees

Paise Only

by cash being the amount of

TOPIC APPROVAL FORM FEE PHD

TOPIC APPROVAL FEES Rs. **********200.00

Total: Rs. **********200.00

Cash Receiver's Signature **Finance & Accounts Officer**

1. Receipt subject to the realisation of Cheque / Draft etc., it tendered.
2. No duplicate receipt will be issued.
3. The statement of marks should be collected from Ground Floor, Window No. 2 after 15 working days on production of the receipt and if not collected within 3 months no responsibility lies on the University.